

REMARKS

Summary

Independent Claims 1, 9, and 17 recite at least one feature not disclosed or suggested by the patent to Kagoshima. Therefore, is the outstanding rejection of these claims over this patent still proper?

Status of the claims

Claims 1-19 are pending. Claims 2 and 10 have been amended. Claims 1, 2, 4, 6, 9, 10, 12, 14, and 17-19 are independent.

Requested action

Applicants respectfully request the Examiner to reconsider and withdraw the outstanding rejection in view of the following remarks.

Allowable subject matter

Applicants gratefully acknowledge the allowance of Claims 4-8, 12-16, 18, and 19 and the indication that Claims 2 and 10 are merely objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims. In response, Applicants have redrafted Claims 2 and 10 in independent form, thereby rendering them allowable.

Rejection

Claims 1, 3, 9, 11, and 17 are rejected under 35 U.S.C. § 102(e) as being anticipated by U.S. Patent No. 5,890,118 (Kagoshima, et al.).

Response to rejection

This rejection is respectfully traversed for the following reasons.

Independent Claim 1 relates to a speech synthesis apparatus for performing speech synthesis by using pitch marks, comprising first calculation means for calculating a distance between first two pitch marks of a voiced portion of speech data to be processed, second calculation means for calculating a difference between adjacent inter-pitch-mark distances, and management means for storing the calculation results obtained by the first and second calculation means in a file and managing the results.

By this arrangement, the size of the file used to manage pitch marks of speech can be reduced.

In contrast, the patent to Kagoshima is designed to provide a speech synthesis apparatus capable of producing synthetic speech excellent in naturalness by reducing the discontinuity at the boundary between frames, as discussed at column 2, lines 24-27. To achieve this function, this patent discloses a device that interpolates between typical waveforms to obtain a plurality of interpolation signals each having at least one of an interpolation pitch period and a signal level that changes smoothly between corresponding frames, as recited in Claim 1 of this patent. But the Kagoshima patent is not understood to disclose or suggest first calculation means for calculating a distance between first two *pitch marks* of a voiced portion of speech data to be processed, as recited by Claim 1. Rather, this patent is understood to merely interpolate *pitch*

periods corresponding to two consecutive sound frames. Therefore, this patent also does not disclose or suggest second calculation means for calculating a difference between adjacent inter-pitch-mark distances, or management means for storing the calculation results obtained by the first and second calculation means in a file and managing the results. Since the Kagoshima patent fails to disclose these three features of Claim 1, Claim 1 is allowable over this patent. And since independent Claims 9 and 17 are corresponding method and program-code claims, they are allowable for similar reasons.

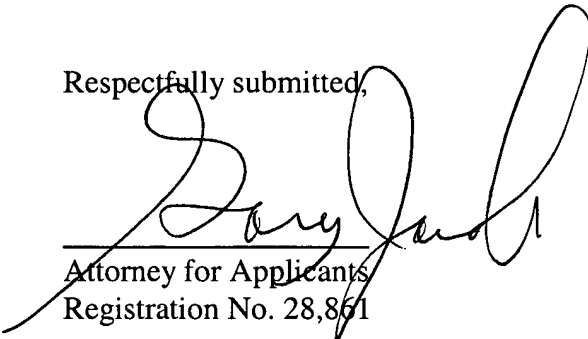
The dependent claims are allowable for the reasons given with respect to the independent claims and because they recite features which are patentable in their own right. Individual consideration of the dependent claims is respectfully solicited.

In view of the above amendments and remarks, the claims are now in allowable form. Therefore, early passage to issue is respectfully solicited.

A

Applicants' undersigned attorney may be reached in our Washington, D.C. office by telephone at (202) 530-1010. All correspondence should continue to be directed to our below-listed address.

Respectfully submitted,



Attorney for Applicants
Registration No. 28,861

FITZPATRICK, CELLA, HARPER & SCINTO
30 Rockefeller Plaza
New York, New York 10112-3801
Facsimile No.: (212) 218-2200
GMJ/cm

DC_MAIN 87595 v 1

VERSION WITH MARKINGS TO SHOW CHANGES MADE TO THE SPECIFICATION

Please substitute the following paragraph for the paragraph starting at page 1, line 16 and ending at line 18.

--In this scheme, information (pitch mark) about the position of each pitch must be recorded concurrently with the storage of speech waveform data.--

Please substitute the following paragraph for the paragraph starting at page 2, line 10 and ending at line 12.

--first calculation means for calculating [a] the distance between first two pitch marks of a voiced portion of speech data to be processed;--

Please substitute the following paragraph for the paragraph starting at page 2, line 13 and ending at line 15.

--second calculation means for calculating [a] the difference between adjacent inter-pitch-mark distances; and--

Please substitute the following paragraph for the paragraph starting at page 3, line 7 and ending at line 10.

--subtraction means for subtracting the maximum value d_{max} or the minimum value d_{min} from the length d on the basis of the comparison results obtained by the first and second comparison means; and--

Please substitute the following paragraph for the paragraph starting at page 3, line 22 and ending at line 25.

--storage means for storing a file for, managing [a] the distance between first two pitch marks of a voiced portion of speech data to be processed and [a] the difference between adjacent inter-pitch-mark distances;--

Please substitute the following paragraph for the paragraph starting at page 4, line 4 and ending at line 8.

--calculation means for calculating [a] the next pitch mark position from a pitch mark position calculated immediately before the calculation, [a] the pitch mark distance to an adjacent pitch mark, and the distance and difference loaded by the first and second loading means.--

Please substitute the following paragraph for the paragraph starting at page 4, line 15 and ending at line 17.

--[the] a first calculation step of calculating [a] the distance between first two pitch marks of a voiced portion of speech data to be processed;--

Please substitute the following paragraph for the paragraph starting at page 4, line 18 and ending at line 20.

--[the] a second calculation step of calculating [a] the difference between adjacent inter-pitch-mark distances; and--

Please substitute the following paragraph for the paragraph starting at page 4, line 21 and ending at line 23.

--[the] a management step of storing the calculation results obtained in the first and second calculation steps in a file and managing the results.--

Please substitute the following paragraph for the paragraph starting at page 5, line 4 and ending at line 8.

--[the] a first comparison step of, when [a] the length of speech data to be processed is represented by d , and [a] the maximum value d_{max} and a minimum value d_{min} are defined for a predetermined word length, comparing the length d with the maximum value d_{max} ;--

Please substitute the following paragraph for the paragraph starting at page 5, line 9 and ending at line 11.

--[the] a second comparison step of comparing the length d with the minimum value d_{min} on the basis of the comparison result obtained in the first comparing step;--

Please substitute the following paragraph for the paragraph starting at page 5, line 12 and ending at line 15.

--[the] a subtraction step of subtracting the maximum value d_{max} or the minimum value d_{min} from the length d on the basis of the comparison results obtained in the first and second comparison steps; and--

Please substitute the following paragraph for the paragraph starting at page 5, line 16 and ending at line 20.

--[the] a management step of storing the difference obtained in the subtraction step or the length d in the file and managing the difference or the length on the basis of the comparison results obtained in the first and second comparison steps.--

Please substitute the following paragraph for the paragraph starting at page 6, line 1 and ending at line 4.

--[the] a storage step of storing a file for managing [a] the distance between the first two pitch marks of a voiced portion of speech data to be processed and [a] the difference between adjacent inter-pitch-mark distances;--

Please substitute the following paragraph for the paragraph starting at page 6, line 5 and ending at line 6.

--[the] a first loading step of loading the distance between the first two pitch marks of the voiced portion;--

Please substitute the following paragraph for the paragraph starting at page 6, line 7 and ending at line 8.

--[the] a second loading step of loading the difference between the adjacent inter-pitch-mark distances; and--

Please substitute the following paragraph for the paragraph starting at page 6, line 9 and ending at line 13.

--[the] a calculation step of calculating a next pitch mark position from a pitch mark position calculated immediately before the calculation, a pitch mark distance to an adjacent pitch mark, and the distance and difference loaded in the first and second loading steps.--

Please substitute the following paragraph for the paragraph starting at page 6, line 21 and ending at line 23.

--a program code for the first calculation step of calculating [a] the distance between the first two pitch marks of a voiced portion of speech data to be processed;--

Please substitute the following paragraph for the paragraph starting at page 6, line 24 and ending at line 26.

--a program code for the second calculation step of calculating [a] the difference between adjacent inter-pitch-mark distances; and--

Please substitute the following paragraph for the paragraph starting at page 7, line 11 and ending at line 15.

--a program code for the first comparison step of, when [a] the length of speech data to be processed is represented by d, and [a] the maximum value dmax and a minimum value dmin are defined for a predetermined word length, comparing the length d with the maximum value dmax;--

Please substitute the following paragraph for the paragraph starting at page 7, line 20 and ending at line 23.

--a program code for the subtraction step of subtracting the maximum value d_{max} or the minimum value d_{min} from the length d on the basis of the comparison results obtained in the first and second comparison steps; and--

Please substitute the following paragraph for the paragraph starting at page 8, line 10 and ending at line 14.

--a program code for the storage step of storing a file for managing [a] the distance between the first two pitch marks of a voiced portion of speech data to be processed and [a] the difference between adjacent inter-pitch-mark distances;--

Please substitute the following paragraph for the paragraph starting at page 9, line 11 and ending at line 13.

--Fig. 2 is a flow chart showing [pitch mark data file generation] pitch-mark-data, file-generation processing executed in the first embodiment of the present invention;--

Please substitute the following paragraph for the paragraph starting at page 9, line 23 and ending at line 25.

--Fig. 6 is a flow chart showing [pitch mark data file loading] pitch-mark-data, file-loading processing executed in the second embodiment of the present invention; and--

Please substitute the following paragraph for the paragraph starting at page 10, line 10 and ending at line 24.

--Reference numeral 103 denotes a CPU for performing numerical operation/control, control on the respective components of the apparatus, and the like, which are executed in the present invention; 102[,] denotes a RAM serving as a work area for processing executed in the present invention, a temporary saving area for various data and having an area for storing a [pitch mark data] pitch-mark-data file 101a; 101[,] denotes a ROM storing various control programs such as programs executed in the present invention, for managing [pitch mark] pitch-mark data used for speech synthesis; 109[,] denotes an external storage unit serving as an area for storing processed data; and 105[,] denotes a D/A converter for converting the digital speech data synthesized by the speech synthesis apparatus into analog speech data and outputting it from a loudspeaker 110.--

Please substitute the following paragraph for the paragraph starting at page 10, line 25 and ending at page 11, line 9.

--Reference numeral 106 denotes a display control unit for controlling a display 111 when the processing state and processing results of the speech synthesis apparatus, and a user interface are to be displayed; 107[,] denotes an input control unit for recognizing key information input from a keyboard 112 and executing the designated processing; 108[,] denotes a communication control unit for controlling transmission/reception of data through a communication network

113; and 104[,] denotes a bus for connecting the respective components of the speech synthesis apparatus to each other.--

Please substitute the following paragraph for the paragraph starting at page 11, line 10 and ending at line 12.

--[Pitch mark data file generation] Pitch-mark-data, file-generation processing executed in the first embodiment will be described next with reference to Fig. 2.--

Please substitute the following paragraph for the paragraph starting at page 11, line 13 and ending at line 15.

--Fig. 2 is a flow chart showing [pitch mark data file generation] pitch-mark-data, file generation processing executed in the first embodiment of the present invention.--

Please substitute the following paragraph for the paragraph starting at page 12, line 18 and ending at page 13, line 3.

--If it is determined that the voiced portion ends (YES in step S6), the flow advances to step S10 to record a voiced portion end signal indicating the end of the voiced portion in the [pitch mark data] pitch-mark-data file 101a. Note that any signal can be used as the voiced portion end signal as long as it can be discriminated from an inter-pitch-mark distance. In step S11, it is checked whether the speech data has ended. If it is determined that the speech data has

not ended (NO in step S11), the flow advances to step S12. If it is determined that the speech data has ended (YES in step S11), the processing is terminated.--

Please substitute the following paragraph for the paragraph starting at page 13, line 4 and ending at line 16.

--It is determined in step S1 that the first segment of the speech data is an unvoiced portion (NO in step S1), the flow advances to step S3 to record unvoiced portion start information indicating that "the first segment is an unvoiced portion" in the pitch mark data file 101a. In step S12, [a] the distance d_s between the voiced portion and the next voiced portion (i.e., the length of the unvoiced portion) is recorded in the pitch mark data file 101a. In step S13, it is checked whether the speech data has ended. If it is determined that the speech data has not ended (NO in step S13), the flow advances to step S4. If it is determined that the speech data has ended (YES in step S13), the processing is terminated.--

Please substitute the following paragraph for the paragraph starting at page 13, line 17 and ending at line 22.

--As described above, according to the first embodiment, since the respective pitch marks in each voiced portion are managed by using the distances between the adjacent pitch marks, all the pitch marks in each voiced portion need not be managed. This can reduce the size of the [pitch mark data] pitch-mark-data file 101a.--

Please substitute the following paragraph for the paragraph starting at page 13, line 23 and ending at page 14, line 4.

--In the first embodiment, step S10 may be replaced with step S14 of counting the number (n) of pitch marks in each voiced portion and step S15 of recording the counted number n of pitch marks in the [pitch mark data] pitch-mark-data file 101a, as shown in Fig. 4. In this case, the processing in step S6 amounts to checking whether the value of the loop counter i is equal to the number n of pitch marks.--

Please substitute the following paragraph for the paragraph starting at page 14, line 15 and ending at line 21.

--First of all, in step S16, d is compared with dmax. If d is equal to or larger than dmax (YES in step S16), the flow advances to step S17 to record the maximum value dmax in the [pitch mark data] pitch-mark-data file 101a. In step S18, dmax is subtracted from d, and the flow returns to step S16. If it is determined that d is smaller than dmax (NO in step S16), the flow advances to step S19.--

Please substitute the following paragraph for the paragraph starting at page 15, line 8 and ending at line 11.

--In the second embodiment, [pitch mark data file] pitch-mark-data-file loading processing of loading data from the [pitch mark data] pitch-mark-data file 101a recorded in the first embodiment will be described with reference to Fig. 6.--

Please substitute the following paragraph for the paragraph starting at page 15, line 12 and ending at line 14.

--Fig. 6 is a flow chart showing [pitch mark data file] pitch-mark-data-file loading processing executed in the second embodiment of the present invention.--

Please substitute the following paragraph for the paragraph starting at page 15, line 15 and ending at line 26.

--First of all, in step S23, start information indicating whether the start of speech data to be processed is a voice or unvoiced portion, is loaded from a [pitch mark data] pitch-mark-data file 101a. It is then checked in step S24 whether the loaded start information is voiced portion start information. If voiced portion start information is determined (YES in step S24), the flow advances to step S25 to load a first inter-pitch-mark distance (distance between a first pitch mark p_1 and a second pitch mark p_2 of the voiced portion) d_1 from the pitch mark data file 101a. Note that the second pitch mark p_2 is located at $p_1 + d_1$.--

Please substitute the following paragraph for the paragraph starting at page 16, line 1 and ending at line 11.

--In step S26, the value of a loop counter i is initialized to 2. In step S27, a difference d_i (data corresponding the length of one word) from the [pitch mark data] pitch-mark-data file 101a. In step S28, it is checked whether the loaded difference d_i is a voiced portion end signal. If it is determined that the difference is not a voiced portion end signal (NO in step S28), the flow

advances to step S29 to calculate [a] the next inter-pitch-mark distance d_i and the pitch mark position p_{i+1} from a pitch mark position p_i , the inter-pitch-mark distance d_{i-1} , and d_r obtained in the past.--

Please substitute the following paragraph for the paragraph starting at page 16, line 11 and ending at line 17.

--The following equations can be formulated from p_i , d_{i-1} , d_r , d_i , and p_{i+1} . The next inter-pitch-mark distance d_i and the pitch mark position p_{i+1} , can be calculated by using these equations.

$$d_i = d_{i-1} + d_r \quad \dots (1)$$

$$p_{i+1} = p_i + d_i \quad \dots (2)--$$

Please substitute the following paragraph for the paragraph starting at page 17, line 11 and ending at line 16.

--As described above, according to the second embodiment, since pitch marks can be loaded by using the [pitch mark data] pitch-mark-data file 101a managed by the processing described in the first embodiment, the size of data to be processed decreases to improve [the] processing efficiency.--

Please substitute the following paragraph for the paragraph starting at page 17, line 24 and ending at page 18, line 2.

--Assume that the [data length] data-length information of loaded speech data is stored in a register d, and a maximum value dmax (e.g., 127), a minimum value dmin (e.g., -127), and a voiced portion end signal are defined for a given word length (e.g., 8. bits) in Fig. 5.--

Please substitute the following paragraph for the paragraph starting at page 18, line 3 and ending at line 12.

--First of all, in step S34, the register d is initialized to 0. In step S35, the data d_r corresponding to the length of one word is loaded from the [pitch mark data] pitch-mark-data file 101a. It is then checked in step S36 whether d_r is a voiced portion end signal. If it is determined that the d_r is a voiced portion end signal (YES in step S36), the processing is terminated. If it is determined that d_r is not a voiced portion end signal (NO in step S36), the flow advances to step S37 to add d_r to the contents of the register d.--

Please substitute the following paragraph for the paragraph starting at page 18, line 18 and ending at line 23.

--Note that the present invention may be applied to either a system constituted by a plurality of pieces of equipments (e.g., a host computer, an interface device, a reader, a printer, and the like), or an apparatus consisting of a single piece of equipment (e.g., a copying machine, a facsimile apparatus, or the like).--

VERSION WITH MARKINGS TO SHOW CHANGES MADE TO THE CLAIMS

2. (Amended) [The] A speech synthesis apparatus for performing speech synthesis by using pitch marks, comprising:

first calculation means for calculating a distance between first two pitch marks of a voiced portion of speech data to be processed;

second calculation means for calculating a difference between adjacent inter-pitch-mark distances; and

management means for storing the calculation results obtained by said first and second calculation means in a file and managing the results [according to claim 1],

wherein said management means further calculates an inter-voiced-portion distance as a distance between voiced portions on both sides of an unvoiced portion, stores the distance in the file, and manages the distance.

10. (Amended) [The] A control method for a speech synthesis apparatus for performing speech synthesis by using pitch marks, comprising:

a first calculation step of calculating a distance between first two pitch marks of a voiced portion of speech data to be processed;

a second calculation step of calculating a difference between adjacent inter-pitch-mark distances; and

a management step of storing the calculation results obtained in said first and second calculation steps in a file and managing the results [according to claim 9],

wherein said management step further comprises calculating an inter-voiced-portion distance as a distance between voiced portions on both sides of an unvoiced portion, storing the distance in the file, and managing the distance.

DC_MAIN 87957 v 1